GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 21:36:48 ; Search time 243 Seconds
(without alignments)
4707.881 Million cell updates/sec

Title: US-09-914-324A-3
Perfect score: 508
Sequence: 1 cccaaaatggcggcagcgat.....aaagtccagttggattctgg 508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	508	100.0	508	21	AAA74978	DNA encoding a hum
c	2	452.8	89.1	4476	23	ABV25615	Human prostate exp
	3	447.2	88.0	482	21	AAC03896	Human secreted pro
	4	415	81.7	3208	22	AAD12859	Human novel protei
	5	333	65.6	1482	21	AAC75216	Human ORFX ORF771
	6	327	64.4	327	21	AAA96882	Nucleotide sequenc
	7	323	63.6	504	21	AAA74980	DNA encoding a mur
c	8	321	63.2	398	23	AAS86841	DNA encoding novel
	9	319.8	63.0	539	22	AAH97860	Murine 7-transmemb
	10	319.8	63.0	586	22	AAH97862	Murine 7-transmemb
c	11	301.2	59.3	5347	24	AAS94844	Human DNA sequence
	12	286	56.3	402	21	AAC06296	Human secreted pro
	13	283.4	55.8	557	22	AAH97861	Murine 7-transmemb
c	14	272.8	53.7	418	24	ABL64342	Stomach cancer rel
c	15	271	53.3	416	23	AAS86840	DNA encoding novel
c	16	265.2	52.2	812	24	ABQ54871	Human ovarian anti
	17	221.4	43.6	254	16	AAT26036	Human gene signatu
	18	219.6	43.2	485	21	AAA43288	Xenopus secreted e
	19	209	41.1	509	23	AAS86845	DNA encoding novel
	20	198.2	39.0	692	23	ABL22527	Drosophila melanog
	21	191	37.6	202	23	AAS58790	cDNA #1466 encodin
c	22	185.4	36.5	4099	23	AAS86847	DNA encoding novel
	23	176.4	34.7	11301	22	AAK71591	Human immune/haema
	24	174.2	34.3	635	21	AAC39854	Arabidopsis thalia
	25	173	34.1	390	24	ABQ85150	Arabidopsis thalia
	26	153.2	30.2	411	21	AAC57009	Pinus radiata tran
	27	152.2	30.0	2841	23	ABL22526	Drosophila melanog
	28	150.6	29.6	666	23	AAS86844	DNA encoding novel
	29	143.8	28.3	346	22	AAL11142	Human breast cance
	30	143.8	28.3	353	22	AAL18905	Human breast cance
	31	138.2	27.2	560	23	ABL18585	Drosophila melanog
	32	138.2	27.2	2560	23	ABL18584	Drosophila melanog
	33	138.2	27.2	3038	23	ABL02194	Drosophila melanog
	34	138.2	27.2	4760	23	ABL20826	Drosophila melanog
	35	138.2	27.2	4764	23	ABL11922	Drosophila melanog
	36	134.4	26.5	358	21	AAA31220	Plant microsatelli
c	37	131	25.8	358	23	AAS86842	DNA encoding novel
c	38	131	25.8	386	23	AAS86843	DNA encoding novel
	39	126.2	24.8	480	21	AAA74979	DNA encoding a yea
	40	111.8	22.0	345	21	AAA31103	Plant microsatelli
	41	99.6	19.6	416	22	AAI89292	Human polynucleoti
	42	90.6	17.8	1140	20	AAX87313	Mouse sensitive to
	43	76.4	15.0	271	24	ABL71361	Corn tassel-derive
	44	76	15.0	754	20	AAX87331	Human sensitive to
	45	76	15.0	754	20	AAX87314	Human sensitive to

ALIGNMENTS

RESULT 1
AAA74978
ID AAA74978 standard; DNA; 508 BP.
XX
AC AAA74978;
XX
DT 02-JAN-2001 (first entry)
XX
DE DNA encoding a human cullin-interacting RING-H2 finger protein (Rbx1).
XX
KW Cullin-interacting RING-H2 finger protein; Ring box protein; Rbx1;
KW tumour suppressor; carcinoma; Ring box associated carcinoma;
KW von Hippel-Lindau complex; ubiquitin conjugation; renal carcinoma;
KW cerebellar hemangioblastoma; hemangioma; retinal angiomata;
KW pheochromocytomas; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7..333
FT /*tag= a
FT /product= "cullin-interacting RING-H2 finger protein
FT (Rbx1)"
XX
PN WO200050445-A1.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04838.
XX
PR 26-FEB-1999; 99US-0121787.
XX

PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Conaway JA, Conaway RC, Kamura T;
XX
DR WPI; 2000-572067/53.
DR P-PSDB; AAB08813.
XX
PT Cullin interacting RING-H2 finger protein, a component of von
PT Hippel-Lindau tumour suppressor complex and Skp1-Cdc53p-F-box protein
PT (SCF) ubiquitin ligase, useful for diagnosing and treating Ring box
PT protein associated carcinomas -
XX
PS Claim 3; Page 35; 37pp; English.
XX
CC The present sequence encodes a human cullin-interacting RING-H2 finger
CC protein (Ring box protein), designated Rbx1. The polypeptide is a tumour
CC suppressor. Rbx1 is useful for diagnosing a predisposition of a patient
CC to certain carcinomas. It is also useful for treating Ring box protein
CC associated carcinomas or augmenting metabolically deficient system in
CC animals. Rbx1 is also useful for evaluating the effectiveness of a
CC therapeutic treatment for Ring box associated carcinomas. Rbx1 can be
CC used to screen for agents which augment or inhibit the activity of
CC other cullin-containing ubiquitin ligase and of the VHL (von Hippel-
CC Lindau) complex controlling the conjugation of ubiquitin or ubiquitin-
CC like proteins to various sets of target proteins. Carcinomas which may
CC be treated include renal carcinomas, cerebellar hemangioblastomas and
CC hemangiomas, retinal angiomata and pheochromocytomas.
XX
SQ Sequence 508 BP; 126 A; 106 C; 124 G; 152 T; 0 other;

Query Match 100.0%; Score 508; DB 21; Length 508;
Best Local Similarity 100.0%; Pred. No. 5.4e-136;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CCCAAAATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAAAATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGC 60

Qy      61 AAGAAGCGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAGAAGCGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTG 120

Qy     121 GTTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCT 180

Qy     181 AACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCT 240

Qy     241 TTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAAC 300

Qy     301 AGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAAT 360

Qy     361 TGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAA 420

Qy     421 CTGTGTTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTGTGTTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTG 480

Qy     481 TGTCAAATAAAGTCCAGTTGGATTCTGG 508
          ||||||||||||||||||||||||||||
Db     481 TGTCAAATAAAGTCCAGTTGGATTCTGG 508
```

RESULT 2
ABV25615/c

ID ABV25615 standard; cDNA; 4476 BP.
XX
AC ABV25615;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 25606.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 5119-5120; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 4476 BP; 1085 A; 1129 C; 1015 G; 1247 T; 0 other;

Query Match 89.1%; Score 452.8; DB 23; Length 4476;
Best Local Similarity 95.2%; Pred. No. 9.9e-120;
Matches 497; Conservative 0; Mismatches 2; Indels 23; Gaps 2;

```
Qy       9 GGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAGCG 68
          | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 GCCCGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAGCG 959

Qy      69 CTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAA 128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 CTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAA 899

Qy     129 CTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGC 188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 CTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGC 839

Qy     189 GTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTT 248
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 GTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTT 779

Qy    249 CCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTG 308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 CCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTG 719

Qy    309 GGAATTCCAAAA----------------------GTATGGGCACTAGGAAAAGACTTCTT 346
          ||||||||||||                      ||||||||||||||||||||||||||
Db    718 GGAATTCCAAAAAGTAATTTACTTTGTCTTTCAGGTATGGGCACTAGGAAAAGACTTCTT 659

Qy    347 CCATCAAGCTTAATTGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAATTA 406
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    658 CCATCAAGCTTAATTGTTTTGTTATTCATTTAA-TGACTTTCCCTGCTGTTACCTAATTA 600

Qy    407 CAAATTGGATGGAACTGTGTTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGC 466
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    599 CAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGC 540

Qy    467 CATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCTGG 508
          |||||||||||||||||||||||||||||||||||||||||||
Db    539 CATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCTGG 498
```

RESULT 3
AAC03896
ID AAC03896 standard; cDNA; 482 BP.
XX
AC AAC03896;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3894.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG03890.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures
XX
PS Claim 1; SEQ ID 3894; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 482 BP; 116 A; 103 C; 121 G; 140 T; 2 other;

```
  Query Match             88.0%;  Score 447.2;  DB 21;  Length 482;
  Best Local Similarity   99.3%;  Pred. No. 1.6e-118;
  Matches  457;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      2 CCAAAATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCA 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     24 CCAAAATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCA 83

Qy     62 AGAAGCGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGG 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     84 AGAAGCGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGG 143

Qy    122 TTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTA 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 TTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTA 203

Qy    182 ACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTT 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 ACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTT 263

Qy    242 TTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACA 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    264 TTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACA 323

Qy    302 GAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATT 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324 GAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATT 383

Qy    362 GTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAAC 421
          ||||||||||||||||: ||:||||||||||||||||||||||||||||||||||||||||
Db    384 GTTTTGTTATTCATTTAR-TGMCTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAAC 442

Qy    422 TGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCT 461
          |||||||||||||||||||||||||||||||||||||||||
Db    443 TGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCT 482
```

RESULT 4
AAD12859
ID AAD12859 standard; cDNA; 3208 BP.
XX
AC AAD12859;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human novel protein (NHP) cDNA with flanking sequences.
XX
KW Human; novel protein; NHP; protease; gene therapy; aminopeptidase;
KW obesity; diabetes; high blood pressure; connective tissue disorder;
KW infertility; infectious disease; antidiabetic; antiinfertility;
KW antibacterial; anorectic; antiviral; ss.
XX
OS Homo sapiens.
XX
PN WO200159134-A1.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US04433.
XX
PR 11-FEB-2000; 2000US-0181924.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX
DR WPI; 2001-488976/53.
XX

```
PT   Isolated nucleic acid molecules encoding proteins with similarity to
PT   aminopeptidases useful in the treatment of diabetes and obesity -
XX
PS   Disclosure; Page 47-48; 51pp; English.
XX
CC   The present sequence is human novel protein (NHP) cDNA with flanking
CC   sequences. NHP share sequence similarity with mammalian proteases.
CC   Proteases cleave protein substrates as part of degradation, maturation
CC   and secretory pathways within the body. NHP cDNA is used in gene therapy.
CC   NHP nucleic acids encode proteins with similarity to aminopeptidases and
CC   are useful in the treatment of diabetes, obesity, high blood pressure,
CC   connective tissue disorder, infertility and infectious disease.
XX
SQ   Sequence 3208 BP; 882 A; 750 C; 772 G; 804 T; 0 other;

  Query Match             81.7%;  Score 415;  DB 22;  Length 3208;
  Best Local Similarity   98.6%;  Pred. No. 6.3e-109;
  Matches  429;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy      74 AAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTG 133
           ||   ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2765 AAAAAAAAAAATGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTG 2824

Qy     134 CCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCG 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2825 CCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCG 2884

Qy     194 CTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACT 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2885 CTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACT 2944

Qy     254 GCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAAT 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2945 GCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAAT 3004

Qy     314 TCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTC 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3005 TCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTC 3064

Qy     374 ATTTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTC 433
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3065 ATTTAA-TGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTC 3123

Qy     434 TGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGT 493
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    3124 TGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGGATTCTGTGTCAAATAAAGT 3183

Qy     494 CCAGTTGGATTCTGG 508
           |||||||||||||||
Db    3184 CCAGTTGGATTCTGG 3198


RESULT 5
AAC75216
ID   AAC75216 standard; cDNA; 1482 BP.
XX
AC   AAC75216;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF771 polynucleotide sequence SEQ ID NO:1541.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB41007.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 1265; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 1482 BP; 308 A; 408 C; 434 G; 332 T; 0 other;

  Query Match             65.6%;  Score 333;  DB 21;  Length 1482;
  Best Local Similarity   99.7%;  Pred. No. 1.7e-85;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy     164 GCATAGAATGTCAAGCTAACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGG 223
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1130 GCATAGAATGTCAAGCTAACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGG 1189

Qy     224 GAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGG 283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1190 GAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGG 1249

Qy     284 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTT 343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1250 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTT 1309

Qy     344 CTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAA 403
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    1310 CTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAA-TGACTTTCCCTGCTGTTACCTAA 1368
```

```
Qy    404 TTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGT 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369 TTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGT 1428

Qy    464 AGCCATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCTGG 508
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1429 AGCCATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCTGG 1473


RESULT 6
AAA96882
ID   AAA96882 standard; DNA; 327 BP.
XX
AC   AAA96882;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Nucleotide sequence of human ring finger protein ROC1.
XX
KW   ROC1; ROC2; cullin; ring finger protein; APC11; APC complex; SCF pathway;
KW   cullin dependent ubiquitin ligase; CDK inhibitor Sic1 degradation;
KW   tumour; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..327
FT                   /*tag=  a
FT                   /product= "ROC1"
XX
PN   WO200058472-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08592.
XX
PR   31-MAR-1999;   99US-0127261.
PR   22-NOV-1999;   99US-0166927.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Xiong Y,  Ohta T;
XX
DR   WPI; 2000-647235/62.
DR   P-PSDB; AAB19160.
XX
PT   Novel nucleic acid encoding cullin regulating ring finger proteins,
PT   termed as ROC proteins similar to anaphase-promoting complex 11, for
PT   therapeutic and diagnostic use -
XX
PS   Claim 1; Fig 2A; 83pp; English.
XX
CC   The present sequence encodes a human ROC1 ring finger protein. The
CC   specification also describes human ROC2. ROC1 and ROC2 are similar
CC   to APC11, a subunit of the APC complex. The proteins stimulate cullin
CC   dependent ubiquitin ligase activity. ROC1 functions in vivo as an
CC   essential regulator of CDK inhibitor Sic1 degradation by the SCF
CC   (undefined) pathway. ROC proteins are useful for screening bioactive
CC   agents that interfere with the binding of ROC proteins with cullin
CC   proteins. Pharmaceutical formulations comprising ROC proteins are
CC   useful for diagnostic and therapeutic purposes, preferably for
CC   diagnosing and treating tumours.
XX
SQ   Sequence 327 BP; 85 A; 75 C; 94 G; 73 T; 0 other;

  Query Match             64.4%;  Score 327;  DB 21;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-84;
  Matches  327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 ATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAG 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAG 60

Qy     67 CGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGAT 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGAT 120

Qy    127 AACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAG 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAG 180

Qy    187 GCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCAC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCAC 240

Qy    247 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAG 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAG 300

Qy    307 TGGGAATTCCAAAAGTATGGGCACTAG 333
          ||||||||||||||||||||||||||
Db    301 TGGGAATTCCAAAAGTATGGGCACTAG 327


RESULT 7
AAA74980
ID   AAA74980 standard; DNA; 504 BP.
XX
AC   AAA74980;
XX
DT   02-JAN-2001  (first entry)
XX
DE   DNA encoding a murine cullin-interacting RING-H2 finger protein (Rbx1).
XX
KW   Cullin-interacting RING-H2 finger protein; Ring box protein; Rbx1;
KW   tumour suppressor; carcinoma; Ring box associated carcinoma;
KW   von Hippel-Lindau complex; ubiquitin conjugation; renal carcinoma;
KW   cerebellar hemangioblastoma; hemangioma; retinal angiomata;
KW   pheochromocytomas; ss.
XX
OS   Mus sp.
XX
PN   WO200050445-A1.
XX
PD   31-AUG-2000.
XX
PF   25-FEB-2000; 2000WO-US04838.
XX
PR   26-FEB-1999;   99US-0121787.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   Conaway JA,  Conaway RC,  Kamura T;
XX
DR   WPI; 2000-572067/53.
XX
PT   Cullin interacting RING-H2 finger protein, a component of von
PT   Hippel-Lindau tumour suppressor complex and Skp1-Cdc53p-F-box protein
PT   (SCF) ubiquitin ligase, useful for diagnosing and treating Ring box
PT   protein associated carcinomas -
XX
PS   Disclosure; Page 35; 37pp; English.
XX
CC   The present sequence encodes a murine cullin-interacting RING-H2 finger
CC   protein (Ring box protein), designated Rbx1. The human Rbx1 polypeptide
CC   is a tumour suppressor. Human Rbx1 is useful for diagnosing a
CC   predisposition of a patient to certain carcinomas. It is also useful
CC   for treating Ring box protein associated carcinomas or augmenting
CC   metabolically deficient system in animals. Human Rbx1 is also useful for
CC   evaluating the effectiveness of a therapeutic treatment for Ring box
CC   associated carcinomas. Human Rbx1 can be used to screen for agents which
CC   augment or inhibit the activity of other cullin-containing ubiquitin
CC   ligase and of the VHL (von Hippel-Lindau) complex controlling the
CC   conjugation of ubiquitin or ubiquitin-like proteins to various sets of
CC   target proteins. Carcinomas which may be treated include renal
```

```
CC   carcinomas, cerebellar hemangioblastomas and hemangiomas, retinal
CC   angiomata and pheochromocytomas.
XX
SQ   Sequence 504 BP; 117 A; 107 C; 137 G; 143 T; 0 other;

  Query Match             63.6%;  Score 323;  DB 21;  Length 504;
  Best Local Similarity   80.6%;  Pred. No. 8.1e-83;
  Matches  394;  Conservative    0;  Mismatches   85;  Indels   10;  Gaps    1;

Qy       2 CCAAAATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCA 61
Db      13 CCAAAATGGCGGCGGCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGGCGCGGGCA 72

Qy      62 AGAAGCGCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGG 121
Db      73 AGAAGCGCTTTGAAGTTAAAAAGTGGAATGCAGTGGCCCTCTGGGCCTGGGACATTGTGG 132

Qy     122 TTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTA 181
Db     133 TTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGTATCGAATGTCAGGCCA 192

Qy     182 ACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTT 241
Db     193 ACCAGGCGTCAGCTACTTCCGAAGAGTGTACGGTTGCATGGGGAGTCTGCAACCATGCTT 252

Qy     242 TTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACA 301
Db     253 TTCATTTCCACTGCATCTCTCGATGGCTCAAAACGAGGCAGGTGTGTCCGTTGGACAACA 312

Qy     302 GAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATT 361
Db     313 GAGAGTGGGAGTTCCAGAAGTATGGGCATTAGGAAAGATTTCCCGCAAGGCGTACCCATC 372

Qy     362 GTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAAC 421
Db     373 TGTTACTTGTCTAGTGA----------CTTCCTGTTAATTATACATTAGATAGAACCATG 422

Qy     422 TGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGT 481
Db     423 GTCCTTTTTCGTTCCTTTGTTTTTGGAGTTTGGTGTTCCCGCAGCCATATTGTATTTTGT 482

Qy     482 GTCAAATAA 490
Db     483 GTAAATAAA 491


RESULT 8
AAS86841/c
ID   AAS86841 standard; cDNA; 398 BP.
XX
AC   AAS86841;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22645.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22654.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 22645; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 398 BP; 122 A; 89 C; 80 G; 107 T; 0 other;

  Query Match             63.2%;  Score 321;  DB 23;  Length 398;
  Best Local Similarity   99.4%;  Pred. No. 2.7e-82;
  Matches  343;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy     164 GCATAGAATGTCAAGCTAACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGG 223
Db     350 GCATAGAATGTCAAGCTAACCA-GCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGG 292

Qy     224 GAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGG 283
Db     291 GAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGG 232

Qy     284 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTT 343
Db     231 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTT 172

Qy     344 CTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAA 403
Db     171 CTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAA-TGACTTTCCCTGCTGTTACCTAA 113

Qy     404 TTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGT 463
Db     112 TTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCTGT 53

Qy     464 AGCCATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCTGG 508
Db      52 AGCCATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCTGG 8


RESULT 9
AAH97860
ID   AAH97860 standard; DNA; 539 BP.
XX
AC   AAH97860;
XX
DT   10-OCT-2001  (first entry)
XX
DE   Murine 7-transmembrane G-protein coupled receptor coding sequence #104.
```

```
XX
KW   Murine; hematopoietic stem cell; signalling; vaccine; 7TM-GPCR;
KW   7-transmembrane G-protein coupled protein receptor; ds.
XX
OS   Mus sp.
XX
PN   WO200160999-A1.
XX
PD   23-AUG-2001.
XX
PF   14-FEB-2001; 2001WO-US04700.
XX
PR   14-FEB-2000; 2000US-0182377.
XX
PA   (IMCL-) IMCLONE SYSTEMS INC.
PA   (UYPR-) UNIV PRINCETON.
XX
PI   Lemischka IR,  Witte L,  Pereira DS;
XX
DR   WPI; 2001-522596/57.
XX
PT   DNA Sequences encoding 7-transmembrane G-protein coupled protein
PT   receptors characteristic of hematopoietic stem cells, useful for
PT   treating leukemia -
XX
PS   Claim 1; Page 62; 176pp; English.
XX
CC   The present invention relates to murine coding sequences for
CC   7-transmembrane G-protein coupled protein receptors (7TM-GPCRs). The
CC   present sequence is one such murine 7TM-GPCR coding sequence. The present
CC   sequence was derived from hematopoietic stem cells. The present sequence
CC   and its corresponding protein are useful in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate 7TM-GPCR expression.
CC   7TM-GPCRs identify specific signalling molecules, to activate an
CC   effector-signalling cascade that triggers an intracellular response and
CC   eventually a biological effect.
XX
SQ   Sequence 539 BP; 128 A; 112 C; 150 G; 149 T; 0 other;

  Query Match           63.0%;  Score 319.8;  DB 22;  Length 539;
  Best Local Similarity  81.0%;  Pred. No. 6.9e-82;
  Matches  388;  Conservative    0;  Mismatches   82;  Indels    9;  Gaps    1;

Qy     16 GCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 75
Db      1 GCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 60

Qy     76 GTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTGCC 135
Db     61 GTTAAAAAGTGGAATGCAGTGGCCCTCTGGGCCTGGGACATTGTGGTTGATAACTGTGCC 120

Qy    136 ATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCGCT 195
Db    121 ATCTGCAGGAACCACATTATGGATCTTTGTATCGAATGTCAGGCCAACCAGGCGTCAGCT 180

Qy    196 ACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACTGC 255
Db    181 ACTTCCGAAGAGTGTACGGTTGCATGGGGAGTCTGCAACCATGCTTTTCATTTCCACTGC 240

Qy    256 ATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAATTC 315
Db    241 ATCTCTCGATGGCTCAAAACGAGGCAGGTGTGTCCGTTGGACAACAGAGAGTGGGAGTTC 300

Qy    316 CAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTCAT 375
Db    301 CAGAAGTATGGGCATTAGGAAAGACTTTCCCGCAAGGCGTACCCATCTGTTACTCGTCTA 360

Qy    376 TTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTCTG 435
Db    361 GTGA---------CTTCCTGTTAATTATACATTAGATAGAACCATGGTCCTTTTTCGTTC 411

Qy    436 CTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGTC 494
Db    412 CTTTGTTTTTGGAGTTTGGTGTTCCCGCAGCCATATTGTATTTTGTGTCAAATAAAGCC 470


RESULT 10
AAH97862
ID   AAH97862 standard; DNA; 586 BP.
XX
AC   AAH97862;
XX
DT   10-OCT-2001  (first entry)
XX
DE   Murine 7-transmembrane G-protein coupled receptor coding sequence #106.
XX
KW   Murine; hematopoietic stem cell; signalling; vaccine; 7TM-GPCR;
KW   7-transmembrane G-protein coupled protein receptor; ds.
XX
OS   Mus sp.
XX
PN   WO200160999-A1.
XX
PD   23-AUG-2001.
XX
PF   14-FEB-2001; 2001WO-US04700.
XX
PR   14-FEB-2000; 2000US-0182377.
XX
PA   (IMCL-) IMCLONE SYSTEMS INC.
PA   (UYPR-) UNIV PRINCETON.
XX
PI   Lemischka IR,  Witte L,  Pereira DS;
XX
DR   WPI; 2001-522596/57.
XX
PT   DNA Sequences encoding 7-transmembrane G-protein coupled protein
PT   receptors characteristic of hematopoietic stem cells, useful for
PT   treating leukemia -
XX
PS   Claim 1; Page 63; 176pp; English.
XX
CC   The present invention relates to murine coding sequences for
CC   7-transmembrane G-protein coupled protein receptors (7TM-GPCRs). The
CC   present sequence is one such murine 7TM-GPCR coding sequence. The present
CC   sequence was derived from hematopoietic stem cells. The present sequence
CC   and its corresponding protein are useful in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate 7TM-GPCR expression.
CC   7TM-GPCRs identify specific signalling molecules, to activate an
CC   effector-signalling cascade that triggers an intracellular response and
CC   eventually a biological effect.
XX
SQ   Sequence 586 BP; 140 A; 122 C; 164 G; 159 T; 1 other;

  Query Match           63.0%;  Score 319.8;  DB 22;  Length 586;
  Best Local Similarity  81.0%;  Pred. No. 7.1e-82;
  Matches  388;  Conservative    0;  Mismatches   82;  Indels    9;  Gaps    1;

Qy     16 GCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 75
Db      1 GCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 60

Qy     76 GTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTGCC 135
Db     61 GTTAAAAAGTGGAATGCAGTGGCCCTCTGGGCCTGGGACATTGTGGTTGATAACTGTGCC 120

Qy    136 ATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCGCT 195
Db    121 ATCTGCAGGAACCACATTATGGATCTTTGTATCGAATGTCAGGCCAACCAGGCGTCAGCT 180

Qy    196 ACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACTGC 255
Db    181 ACTTCCGAAGAGTGTACGGTTGCATGGGGAGTCTGCAACCATGCTTTTCATTTCCACTGC 240

Qy    256 ATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAATTC 315
Db    241 ATCTCTCGATGGCTCAAAACGAGGCAGGTGTGTCCGTTGGACAACAGAGAGTGGGAGTTC 300

Qy    316 CAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTCAT 375
Db    301 CAGAAGTATGGGCATTAGGAAAGACTTTCCCGCAAGGCGTACCCATCTGTTACTCGTCTA 360

Qy    376 TTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTCTG 435
Db    361 GTGA---------CTTCCTGTTAATTATACATTAGATAGAACCATGGTCCTTTTTCGTTC 411

Qy    436 CTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGTC 494
Db    412 CTTTGTTTTTGGAGTTTGGTGTTCCCGCAGCCATATTGTATTTTGTGTCAAATAAAGCC 470


RESULT 10
AAH97862
ID   AAH97862 standard; DNA; 586 BP.
XX
AC   AAH97862;
XX
DT   10-OCT-2001  (first entry)
XX
DE   Murine 7-transmembrane G-protein coupled receptor coding sequence #106.
XX
KW   Murine; hematopoietic stem cell; signalling; vaccine; 7TM-GPCR;
KW   7-transmembrane G-protein coupled protein receptor; ds.
XX
OS   Mus sp.
XX
PN   WO200160999-A1.
XX
PD   23-AUG-2001.
XX
PF   14-FEB-2001; 2001WO-US04700.
XX
PR   14-FEB-2000; 2000US-0182377.
XX
PA   (IMCL-) IMCLONE SYSTEMS INC.
PA   (UYPR-) UNIV PRINCETON.
XX
PI   Lemischka IR,  Witte L,  Pereira DS;
XX
DR   WPI; 2001-522596/57.
XX
PT   DNA Sequences encoding 7-transmembrane G-protein coupled protein
PT   receptors characteristic of hematopoietic stem cells, useful for
PT   treating leukemia -
XX
PS   Claim 1; Page 63; 176pp; English.
XX
CC   The present invention relates to murine coding sequences for
CC   7-transmembrane G-protein coupled protein receptors (7TM-GPCRs). The
CC   present sequence is one such murine 7TM-GPCR coding sequence. The present
CC   sequence was derived from hematopoietic stem cells. The present sequence
CC   and its corresponding protein are useful in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate 7TM-GPCR expression.
CC   7TM-GPCRs identify specific signalling molecules, to activate an
CC   effector-signalling cascade that triggers an intracellular response and
CC   eventually a biological effect.
XX
SQ   Sequence 586 BP; 140 A; 122 C; 164 G; 159 T; 1 other;

  Query Match           63.0%;  Score 319.8;  DB 22;  Length 586;
  Best Local Similarity  81.0%;  Pred. No. 7.1e-82;
  Matches  388;  Conservative    0;  Mismatches   82;  Indels    9;  Gaps    1;

Qy     16 GCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 75
Db      1 GCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 60

Qy     76 GTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTGCC 135
Db     61 GTTAAAAAGTGGAATGCAGTGGCCCTCTGGGCCTGGGACATTGTGGTTGATAACTGTGCC 120

Qy    136 ATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCGCT 195
Db    121 ATCTGCAGGAACCACATTATGGATCTTTGTATCGAATGTCAGGCCAACCAGGCGTCAGCT 180

Qy    196 ACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACTGC 255
Db    181 ACTTCCGAAGAGTGTACGGTTGCATGGGGAGTCTGCAACCATGCTTTTCATTTCCACTGC 240

Qy    256 ATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAATTC 315
Db    241 ATCTCTCGATGGCTCAAAACGAGGCAGGTGTGTCCGTTGGACAACAGAGAGTGGGAGTTC 300
```

```
Db    241 ATCTCTCGATGGCTCAAAACGAGGCAGGTGTGTCCGTTGGACAACAGAGAGTGGGAGTTC 300

Qy    316 CAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTCAT 375
          || ||||||||||| |||||||    |||   ||    |      | | ||  |  ||
Db    301 CAGAAGTATGGGCATTAGGAAAGACTTTCCCGCAAGGCGTACCCATCTGTTACTCGTCTA 360

Qy    376 TTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTCTG 435
           | |         | | ||||||  ||   |  |  | ||   |      |||||   |
Db    361 GTGA---------CTTCCTGTTAATTATACATTAGATAGAACCATGGTCCTTTTTCGTTC 411

Qy    436 CTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGTC 494
          ||||||||||  |||||| |||| | | ||||||||||||||| ||||||||||||| |
Db    412 CTTTGTTTTTGGAGTTTGGTGTTCCCGCAGCCATATTGTATTTTGTGTCAAATAAAGCC 470


RESULT 11
AAS94844/c
ID   AAS94844 standard; DNA; 5347 BP.
XX
AC   AAS94844;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Human DNA sequence #99 expressed during foam cell differentiation.
XX
KW   Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW   cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177389-A2.
XX
PD   18-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US11128.
XX
PR   05-APR-2000; 2000US-195106P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Shiffman D,  Somogyi R,  Lawn R,  Seilhamer JJ,  Porter GJ,  Mikita T;
PI   Tai J;
XX
DR   WPI; 2002-010925/01.
XX
PT   Composition useful for diagnosis of conditions, disorders or diseases
PT   associated with atherosclerosis, comprises several polynucleotides that
PT   are differentially expressed in foam cell development -
XX
PS   Claim 1; Page 151-152; 315pp; English.
XX
CC   The present invention relates to the isolation of human polynucleotide
CC   sequences that are differentially expressed during foam cell
CC   differentiation. The polynucleotide sequences of the invention or a
CC   composition comprising these polynucleotides are useful as a high
CC   throughput method for detecting altered expression of one or more
CC   polynucleotides in a sample. The polynucleotides can be used in the
CC   diagnosis of disorders associated with foam cell development such as
CC   atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC   coronary artery disease. The polynucleotide sequences can also be used
CC   as PCR primers and probes. The polynucleotides of the invention are also
CC   useful in gene therapy. AAS94746-AAS95021 represent the human
CC   polynucleotide sequences of the invention which are differentially
CC   expressed during foam cell differentiation.
XX
SQ   Sequence 5347 BP; 1489 A; 1277 C; 1094 G; 1487 T; 0 other;

  Query Match             59.3%;  Score 301.2;  DB 24;  Length 5347;
  Best Local Similarity   98.7%;  Pred. No. 4e-76;
  Matches  314;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      9 GGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGC-GGGCAAGAAGC 67
          | | ||||||||||||||||||||||||||||||||||||||||||| ||||||||||

Db    519 GCCCGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGGCAAGAAGC 460

Qy     68 GCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATA 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    459 GCTTTGAAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATA 400

Qy    128 ACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGG 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    399 ACTGTGCCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGG 340

Qy    188 CGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACT 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    339 CGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACT 280

Qy    248 TCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGT 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    279 TCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGT 220

Qy    308 GGGAATTCCAAAAGTATG 325
          |||||||||||||||| |
Db    219 GGGAATTCCAAAAGTAGG 202


RESULT 12
AAC06296
ID   AAC06296 standard; cDNA; 402 BP.
XX
AC   AAC06296;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 10371.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 10371; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
```

```
SQ   Sequence 402 BP; 101 A; 86 C; 96 G; 116 T; 3 other;

  Query Match           56.3%;  Score 286;  DB 21;  Length 402;
  Best Local Similarity  99.7%;  Pred. No. 3.2e-72;
  Matches  297;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy    164 GCATAGAATGTCAAGCTAACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGG 223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 GCATAGAATGTCAAGCTAACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGG 165

Qy    224 GAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGG 283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 GAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGG 225

Qy    284 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTT 343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGACTT 285

Qy    344 CTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACCTAA 403
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    286 CTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAA-TGACTTTCCCTGCTGTTACCTAA 344

Qy    404 TTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCT 461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 TTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTCT 402


RESULT 13
AAH97861
ID   AAH97861 standard; DNA; 557 BP.
XX
AC   AAH97861;
XX
DT   10-OCT-2001  (first entry)
XX
DE   Murine 7-transmembrane G-protein coupled receptor coding sequence #105.
XX
KW   Murine; hematopoietic stem cell; signalling; vaccine; 7TM-GPCR;
KW   7-transmembrane G-protein coupled protein receptor; ds.
XX
OS   Mus sp.
XX
PN   WO200160999-A1.
XX
PD   23-AUG-2001.
XX
PF   14-FEB-2001; 2001WO-US04700.
XX
PR   14-FEB-2000; 2000US-0182377.
XX
PA   (IMCL-) IMCLONE SYSTEMS INC.
PA   (UYPR-) UNIV PRINCETON.
XX
PI   Lemischka IR,  Witte L,  Pereira DS;
XX
DR   WPI; 2001-522596/57.
XX
PT   DNA Sequences encoding 7-transmembrane G-protein coupled protein
PT   receptors characteristic of hematopoietic stem cells, useful for
PT   treating leukemia -
XX
PS   Claim 1; Page 62; 176pp; English.
XX
CC   The present invention relates to murine coding sequences for
CC   7-transmembrane G-protein coupled protein receptors (7TM-GPCRs). The
CC   present sequence is one such murine 7TM-GPCR coding sequence. The present
CC   sequence was derived from hematopoietic stem cells. The present sequence
CC   and its corresponding protein are useful in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate 7TM-GPCR expression.
CC   7TM-GPCRs identify specific signalling molecules, to activate an
CC   effector-signalling cascade that triggers an intracellular response and
CC   eventually a biological effect.
XX

SQ   Sequence 557 BP; 135 A; 115 C; 154 G; 153 T; 0 other;

  Query Match           55.8%;  Score 283.4;  DB 22;  Length 557;
  Best Local Similarity  78.3%;  Pred. No. 2e-71;
  Matches  375;  Conservative    0;  Mismatches   81;  Indels   23;  Gaps    2;

Qy     16 GCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 75
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1 GCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGGCGCGGGCAAGAAGCGCTTTGAA 60

Qy     76 GTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTGCC 135
          || |||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db     61 GTTAAAAAGTGGAATGCAGTGGCCCTCTGGGCCTGGGACATTGTGGTTGATAACTGTGCC 120

Qy    136 ATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCGCT 195
          |||||||||||||||||||||||||| || |||||||| || |||||
Db    121 ATCTGCAGGAACCACATTATGGATCTTTGTATCGAATGTCAGGCCAACCA---------- 170

Qy    196 ACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACTGC 255
              | |||||||||| || |||||||||||||| ||||||||||||| |||||||||
Db    171 ----CCGAAGAGTGTACGGTTGCATGGGGAGTCTGCAACCATGCTTTTCATTTCCACTGC 226

Qy    256 ATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAATTC 315
          |||||||| |||||||||||  | |||||||||| |||||||||||||||||||| |||
Db    227 ATCTCTCGATGGCTCAAAACGAGGCAGGTGTGTCCGTTGGACAACAGAGAGTGGGAGTTC 286

Qy    316 CAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTCAT 375
          || |||||||||| |||||||    |||   ||    |     | | ||   |   ||
Db    287 CAGAAGTATGGGCATTAGGAAAGACTTTCCCGCAAGGCGTACCCATCTGTTACTCGTCTA 346

Qy    376 TTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTCTG 435
          | |         | | ||||||  ||   |  |  | ||   |      ||||   |
Db    347 GTGA---------CTTCCTGTTAATTATACATTAGATAGAACCATGGTCCTTTTTCGTTC 397

Qy    436 CTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGTC 494
          ||||||||||  |||||| |||| | | |||||||||||||| ||||||||||||| |
Db    398 CTTTGTTTTTGGAGTTTGGTGTTCCCGCAGCCATATTGTATTTTGTGTCAAATAAAGCC 456


RESULT 14
ABL64342/c
ID   ABL64342 standard; DNA; 418 BP.
XX
AC   ABL64342;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Stomach cancer related gene sequence SEQ ID NO:2679.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
```

```
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   25-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   26-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   27-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236034P.
PR   28-SEP-2000; 2000US-236109P.
PR   28-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-236891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   02-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a
PT   chemical agent to be tested for anti-neoplastic activity, and
PT   determining a change in expression of a gene of a signature gene set -
XX
PS   Claim 1; SEQ ID 2679; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 418 BP; 129 A; 82 C; 89 G; 118 T; 0 other;

  Query Match           53.7%;  Score 272.8;  DB 24;  Length 418;
  Best Local Similarity  86.0%;  Pred. No. 2e-68;
  Matches  349;  Conservative    0;  Mismatches   52;  Indels    5;  Gaps    4;

Qy    101 TCTGGGCCTGGGATATTGTGGTTGATAACTGTGCCATCTGCAGGAACCACATTATGGATC 160
          ||||||| ||||||| |||||||| |||||||||||||||||||||| ||||  |||||||
Db    417 TCTGGGCTTGGGATGTTGTGGTTAATAACTGTGCCATCTGCAGGAATCACAGCATGGATC 358

Qy    161 TTTGCATAGAATGTCAAGCTAACCAGGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCAT 220
          |||||  |||||||||||||||||| | ||| || |||||||||| |||||| || |||
Db    357 --TGCATTGAATGTCAAGCTAACCAAGAGTCTGCCACTTCAGAAGTGTGTACCGTTGCAC 300

Qy    221 GGGGAGTCTGTAACCATGCTTTTCACTTCCACTGCATCTCTCGCTGGCTCAAAACACGAC 280
          |||||| |||||||| |||||||||||| |||||  |||||| |||||||||||||| ||
Db    299 GGGGAGCCTGTAACCGTGCTTTTCACTT-CACTGTCTCTCTCACTGGCTCAAAACACAAC 241

Qy    281 AGGTGTGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCACTAGGAAAAGA 340
          || |||| |  |||||||||||| | | ||||||| ||||||||| ||| || ||||||
Db    240 AGCTGTGCCTGTTGGACAACAGACAATAGGAATTCCCAAAGTATGGACACCAGAAAAAGA 181

Qy    341 CTTCTTCCATCAAGCTTAATTGTTTTGTTATTCATTTAATTGACTTTCCCTGCTGTTACC 400
          |||||||||||||||||| |||||||| ||||||||| |||||| |||| |||| ||
Db    180 ATTCTTCCATCAAGCTTAACTGTTTTGTGATTCATTTAA-TGACTTGCCCTTCTGTCACT 122

Qy    401 TAATTACAAATTGGATGGAACTGTGTTTTTTTCTGCTTTGTTTTTTCAGTTTGCTGTTTC 460
          |||||| ||||| ||| |||||||| | |||||||||||| ||| |  |||||||||
Db    121 TAATTATAAATTAGATAGAACTGTG-TCTTTTCTGCTTTGTCTTTGCCATTTGCTGTTCT 63

Qy    461 TGTAGCCATATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCT 506
          | ||||| ||||||||||||||||||||||||||||||||||||||
Db     62 CGCAGCCACATTGTATTCTGTGTCAAATAAAGTCCAGTTGGATTCT 17


RESULT 15
AAS86840/c
ID   AAS86840 standard; cDNA; 416 BP.
XX
AC   AAS86840;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22644.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22653.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 22644; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
```

```
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 416 BP; 136 A; 83 C; 75 G; 97 T; 25 other;

  Query Match             53.3%;  Score 271;  DB 23;  Length 416;
  Best Local Similarity   83.4%;  Pred. No. 6.6e-68;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels   72;  Gaps    2;

Qy      74 AAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTG 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     416 AAGTGAAAAAGTGGAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGTG 357

Qy     134 CCATCTGCAGGAACCACATTATGGATCTTTGCATAGAATGTCAAGCTAACCAGGCGTCCG 193
           |||||||||||||||||||||||||||||
Db     356 CCATCTGCAGGAACCACATTATGGATCTTT---------------------------- 327

Qy     194 CTACTTCAGAAGAGTGTACTGTCGCATGGGGAGTCTGTAACCATGCTTTTCACTTCCACT 253
                                                 ||||||||||||||||||
Db     326 ----------------------------------------CATGCTTTTCACTTCCACT 308

Qy     254 GCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAAT 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 GCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAGTGGGAAT 248

Qy     314 TCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTC 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     247 TCCAAAAGTATGGGCACTAGGAAAAGACTTCTTCCATCAAGCTTAATTGTTTTGTTATTC 188

Qy     374 ATTTAATTGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTC 433
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 ATTTAA-TGACTTTCCCTGCTGTTACCTAATTACAAATTGGATGGAACTGTGTTTTTTTC 129

Qy     434 TGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGT 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128 TGCTTTGTTTTTTCAGTTTGCTGTTTCTGTAGCCATATTGTATTCTGTGTCAAATAAAGT 69

Qy     494 CCAGTTGGATTCTGG 508
           |||||||||||||||
Db      68 CCAGTTGGATTCTGG 54


Search completed: May 10, 2003, 21:56:21
Job time : 261 secs
```